

A Literature Survey on the Significant Applications of Data Mining in Improving the Quality of Performances of the Systems

¹C. Grace Padma, ² Dr.T. Arunkumar

¹Research Scholar in Ph.D Computer Science, Manonmaniam Sundaranar University, Tirunelveli

²Professor & Research Supervisor

Abstract: *This paper is about the literature survey done on the various significant Applications of Data Mining. Many fields have recognized that application of Data Mining Tools and Techniques have yielded a better result for their useful decision making. Now, since there is a large explosion of data and the data & information are becoming too complex, a better and efficient Data Mining Techniques have to be thought of. Hence, through this survey, one could understand how Data Mining has helped various applications in finding useful knowledges.*

1. APPLICATION OF DATA MINING IN BUSINESS INTELLIGENCE

Typical business intelligence applications of data mining include Risk Analysis (given a set of current customers and their finance/insurance history data, build a predictive model that can be used to classify a new customer into a risk category), Targeted Marketing (given a set of current customers and history on their purchases and their responses to promotions, target new promotions to those most likely to respond), Customer Retention (given a set of past customers and their behavior prior to leaving, predict who is most likely to leave and take proactive action), and Fraud Detection (detect fraudulent activities either proactively or on-line real-time). Many other new application domains are surfacing as we continue to explore and expand upon new data mining opportunities.

2. WORKFLOW QUALITY OF SERVICE MANAGEMENT USING DATA MINING TECHNIQUES

Organizations have been aware of the importance of quality of service (QoS) for competitiveness for some time. It has been widely recognized that workflow systems are a suitable solution for managing the QoS of processes and workflows. The correct management of the QoS of workflows allows for organizations to increase customer satisfaction, reduce internal costs, and increase added value services. Hence, one could design a novel method, composed of several phases, describing how organizations can apply data mining algorithms to predict the QoS for their running workflow instances. A method can be validated using experimentation by applying different data mining algorithms to predict the QoS of workflow.

3. A PROPOSAL FOR THE MANAGEMENT OF MOBILE NETWORK'S QUALITY OF SERVICE (QOS) USING DATA MINING METHODS

Today, the challenge for the service operators is not only to attract and subscribe new users but to retain already subscribed users. To gain a competitive edge over other service operators, the operating personnel have to measure the services provided to their users and the network performance in terms of Quality of Service (QoS) at regular periods. By analyzing the information in these measurements, they can manage the quality of service, which helps to improve their service and network performance. But due to the heavy increase in the number of users in recent years, they find it difficult to elicit essential information from such a large and complex data to manage the QoS using the existing methods. It is here that the recently developed and more powerful data mining methods come in handy. The three data mining methods: Rough Set Theory, Classification and Regression Tree (CART), and Self Organizing Map (SOM) can be used to manage the mobile network QoS.

4. KNOWLEDGE MANAGEMENT IN THE INDUSTRY BASED ON THE USE OF DATA-MINING TECHNIQUES

The current emphasis in theory and practice of knowledge management (KM) is on attempts to understand knowledge creation, transmission, storage and retrieval. Data mining (DM) is intended to provide support in the complex data rich but information poor situations. In the paper we argue that data mining can make a significant contribution to a knowledge management effort. Our goal is to show how data mining techniques can be used for building organizational knowledge, which would lead to a better performance. Finally, one could show how data mining techniques can be used in knowledge management of corporation.

5. DATA MINING CLASSIFICATION TECHNIQUE FOR TALENT MANAGEMENT USING SVM (SUPPORT VECTOR MACHINE)

In Human Resource Management (HRM), the top challenge for HR professionals is managing the organizational talents. The talent management problem can be solved using the classification technique in data mining. There are several classification techniques present such as Decision Tree, Neural Networks, Support vector machine (SVM) and nearest neighbour algorithm. We can also suggest a combined hybrid approach CACC-SVM for potential classification of HR data. This approach yields better accuracy than the traditional classification algorithms because of concise summarization of continuous attributes through CACC discretization and high performing generalized classifier SVM.

6. A MODEL FOR MEASURING E-LEARNING SYSTEMS SUCCESS IN UNIVERSITIES USING DATA MINING

In the era of internet, universities and higher education institutions are increasingly tend to provide e-learning. For suitable planning and more enjoying the benefits of this educational approach, a model for measuring success of e-learning systems is essential. So a survey was also done on the present model used for measuring success of e-learning systems in universities. For this purpose, at first, according to literature review, a conceptual model was designed. Then, based on opinions of 33 experts, and assessing their suggestions, research indicators were finalized. After that, to examine the relationships between components and finalize the proposed model, a case study was done in 5 universities: Amir Kabir University, Tehran University, Shahid Beheshti University, Iran University of Science & Technology and Khaje Nasir Toosi University of Technology. Finally, by analyzing questionnaires completed by 369 instructors, students and alumni, which were e-learning systems user, the final model was designed (MELSS Model)

7. ASSESSING THE IMPACT OF E-LEARNING SYSTEMS ON LEARNERS: A SURVEY STUDY IN THE KSA

With the rapid growth of the use of e-learning systems around the globe, assessing the success and impact of such systems is becoming increasingly important. This paper presents findings from a study of the impact of e-learning systems on university students in the Kingdom of Saudi Arabia. It is asserted that gauging the impact of e-learning systems on learners is central to the development of suitable and effective e-learning systems. Students from two different universities in the KSA were surveyed to capture their perceptions regarding their current e-learning systems. The assessment framework is based on the IS Success/Impact Measurement framework, which has been successfully applied to similar studies on e-Learning, e-Health, and e-Government. Reports on the impacts that the e-learning systems have had on student participant's performance with regard to the depth of learning, customization of learning pace, student productivity, and student satisfaction were possible. The conclusion of this study is that the use of e-learning systems shows a positive impact on student learning.

8. THE ROLE OF QUALITY FACTORS IN INTENTION TO CONTINUE USING AN E-LEARNING SYSTEM IN MALAYSIA

With the advent of the Internet, more and more public universities in Malaysia are putting in effort to introduce e-learning in their respective universities. This entails a significant amount of investments and if the users do not use them the investment would not benefit and would be deemed as a failure. Many previous researches have looked at continued usage as a measure of

A Literature Survey on the Significant Applications of Data Mining in Improving the Quality of Performances of the Systems

success in information system implementation. This research also follows in the same direction. Using a structured questionnaire derived from the literature, data was collected from 1616 undergraduate and post graduate students from public universities in Malaysia. The questionnaire consisted of 3 sections. The first section collected the demographic data, the second section elicited information about information quality, service quality and system quality, section three measured the continuance intention. Data was analyzed using multiple regression analysis. The results indicate that service quality ($\beta = 0.382, p < 0.01$), information quality ($\beta = 0.338, p < 0.01$) and system quality ($\beta = 0.175, p < 0.01$) were positively related to intention to continue usage explaining a total of 59.1% variance. A closer examination reveals that in terms of predictive power service quality had the biggest influence followed by information quality and then system quality. Implications from these findings to e-learning system developers and implementers were found more useful.

9. ORGANIZATIONAL IMPACT OF SYSTEM QUALITY, INFORMATION QUALITY, AND SERVICE QUALITY

Increased organizational dependence on information systems drives management attention towards improving information systems' quality. A recent survey shows that "Improve IT quality" is one of the top concerns facing IT executives. As IT quality is a multidimensional measure, it is important to determine what aspects of IT quality are critical to organizations to help Chief Information Officers (CIOs) to devise effective IT quality improvement strategies. In this research, we model the relationship between information systems' (IS) quality and organizational impact. We hypothesize greater organizational impact in situations in which system quality, information quality and service quality are high. We also hypothesize a positive relationship between system quality and information quality. We test our hypotheses using survey data. Our structural equation model exhibits a good fit with the observed data. Our results show that IS service quality is the most influential variable in this model (followed by information quality and system quality), thus highlighting the importance of IS service quality for organizational performance. This paper contributes theoretically to IS success models through the system quality-to-information quality and IS quality-to-organizational impact links.

10. A SERVICE QUALITY AND SUCCESS MODEL FOR THE INFORMATION SERVICE INDUSTRY WITH THE USAGE OF DATA MINING:

This study proposes and tests a model of library success that shows how information service quality relates to other variables associated with success. If service quality affects success, then it should be possible to compare service quality to other variables believed to affect success. A modified version of the SERVQUAL instrument was evaluated to determine how effectively it measures service quality within the information service industry. Instruments designed to measure information center success and information system successes were evaluated to determine how effectively they measure success in the library system application and how they relate to SERVQUAL.

Responses from 385 end users at two US Army Corps of Engineers libraries were obtained through a mail survey. Results indicate that service quality is best measured with a performance-based version of SERVQUAL, and that measuring importance may be as critical as measuring expectations for management purposes. Results also indicate that service quality is an important factor in success. The findings have implications for the development of new instruments to more effectively measure information service quality and information service success as well as for the development of new models that better show the relationship between information service quality and information service success.

11. CORRELATIONS BETWEEN EXTERNAL KNOWLEDGE AND THE KNOWLEDGE CHAIN AS IMPACTING SERVICE QUALITY

Service Quality continues to be a major concern amongst consumers. As a result, in order to provide the continuous improvement of service quality that will lead to consumer satisfaction, this study explores the impact of external knowledge and knowledge chain on service quality. The results of this study found that absorptive capacity of the external knowledge is indeed an

important source of competitive advantage. Hence, enterprises should apply the knowledge chain in order to gather external knowledge from customers, suppliers and competitors, as well as transforming the knowledge to enhance their service quality.

12. ALTERNATIVE SCALES FOR MEASURING SERVICE QUALITY: A COMPARATIVE ASSESSMENT BASED ON PSYCHOMETRIC AND DIAGNOSTIC CRITERIA

Service quality measurement is an area of growing interest to researchers and managers. It is also an area characterized by debate concerning the need for measuring customer expectations and how they should be measured. Building on a synthesis of the extant literature on customer expectations and service quality measurement, this article identifies unresolved issues and develops three alternative questionnaire formats to address them. It then discusses an empirical study that evaluated the three formats in four different sectors. The article concludes with practical implications and directions for further research stemming from the study's findings.

13. DECISION SUPPORT SYSTEM FOR SERVICE QUALITY MANAGEMENT USING CUSTOMER KNOWLEDGE IN PUBLIC SERVICE ORGANIZATION

As the service quality has been reconsidered in the public sector as well as private enterprises, the need for public sectors to adopt principle and practices of private sectors is concerned with customer-focused approach. However, the different business culture of public service organizations makes it difficult to improve service quality. It is required to establish a structured framework that leads employees to make efforts to improve their service delivery processes and supports continuous improvement of service delivery processes based on the data about the process performance from the customer-perceived value-oriented viewpoint.

This paper proposes a structured framework that identifies the key service processes, validates from customer perspectives and establishes the measurements to monitor based on the data about the process performance. It uses periodic customer satisfaction index (CSI) surveys (S.C. Park) for understanding customer-perceived values. The proposed framework consists of three phases; the questionnaire design, the key process (KP) identification from the integrated viewpoints of importance and contribution, and the key process indicator (KPI) derivation and management. For the application, we established a web-based decision support system for a public service organization for tourism in Korea.

14. THE EFFECTS OF EDUCATION COMPATIBILITY AND TECHNOLOGICAL EXPECTANCY ON E-LEARNING ACCEPTANCE

Discerning what influences a student's acceptance of e-learning is still unclear and has not been well investigated. On the basis of the expectancy-value theory, much effort has been put into identifying the effectual factors regarding the technological expectancy of students. However, aside from technological usage, the adoption of an e-learning system still must consider learning behavior. Thus, researchers should take into consideration both technological and learning expectancies of students while investigating e-learning acceptance. Following mainstream literature on information system acceptance, this study postulates that a student's behavioral intention to accept an e-learning system is determined both by his or her technological expectancy and educational compatibility. Four primary factors, that is, performance expectancy, effort expectancy, social influence, and facilitating conditions, specified in the Unified Theory of Acceptance and Use of Technology (UTAUT) are used to reflect the technological expectancy of students. Further, educational compatibility, which refers the congruence of e-learning systems with the unique leaning expectancies of students, is integrated with the UTAUT to form a new theoretical model for e-learning acceptance. An empirical survey is conducted to examine the proposed model. A total of 626 valid samples were collected from the users of an e-learning system. The findings show that both technological expectancy and educational compatibility are important determinants of e-learning acceptance. However, educational compatibility reveals a greater total effect on e-learning acceptance than does technological expectancy. Implications and practical guidelines for both e-learning developers and practitioners are subsequently presented.

15. A META-ANALYSIS OF E-LEARNING TECHNOLOGY ACCEPTANCE: THE ROLE OF USER TYPES AND E-LEARNING TECHNOLOGY TYPES

Existing literature in the field of e-learning technology acceptance reflects a significant number of independent studies that primarily investigate the causal relationships proposed by technology acceptance theory, such as the technology acceptance model (TAM). To synthesize the existing knowledge in the field of e-learning technology acceptance, we have conducted a systematic literature review of 42 independent papers, mostly published in major journals. Furthermore, in order to view the research context by combining and analyzing the quantitative results of the reviewed research studies, a meta-analysis of the causal effect sizes between common TAM-related relationships was conducted. The main findings of this study, which is the first of its kind, are: (1) TAM is the most-used acceptance theory in e-learning acceptance research, and (2) the size of the causal effects between individual TAM-related factors depends on the type of user and the type of e-learning technology. The results of the meta-analysis demonstrated a moderating effect for user-related factors and technology-related factors for several evaluated causal paths. We have gathered proof that the perceived ease of use and the perceived usefulness tend to be the factors that can influence the attitudes of users toward using an e-learning technology in equal measure for different user types and types of e-learning technology settings.

16. GLOBAL DATA MINING: AN EMPIRICAL STUDY OF CURRENT TRENDS, FUTURE FORECASTS AND TECHNOLOGY DIFFUSIONS

Using a bibliometric approach, this paper analyzes research trends and forecasts of data mining from 1989 to 2009 by locating heading “data mining” in topic in the SSCI database. The bibliometric analytical technique was used to examine the topic in SSCI journals from 1989 to 2009, we found 1181 articles with data mining. This paper implemented and classified data mining articles using the following eight categories—publication year, citation, country/territory, document type, institute name, language, source title and subject area—for different distribution status in order to explore the differences and how data mining technologies have developed in this period and to analyze technology tendencies and forecasts of data mining under the above results. Also, the paper performs the K-S test to check whether the analysis follows Lotka’s law. Besides, the analysis also reviews the historical literatures to come out technology diffusions of data mining. The paper provides a roadmap for future research, abstracts technology trends and forecasts, and facilitates knowledge accumulation so that data mining researchers can save some time since core knowledge will be concentrated in core categories. This implies that the phenomenon “success breeds success” is more common in higher quality publications.

17. EXPLORING AND IMPROVING CLUSTERING BASED STRATEGIES FOR CHEMICAL PROCESS SUPERVISION

In last 12 years, Clustering has received much interest for Process Engineering problems. Particularly, the combination of fuzzy clustering with multivariate statistical techniques for Process Supervision Strategies (PSS) has been studied. The above has led to several approaches. However, some clustering associated problems have been ignored. Also, existing PSS have not been compared. In this work, Clustering based PSS (CPSS) are briefly reviewed and a comparison of it is made. This comparison incorporates some novel strategies that adequately treat some identified problems and it is illustrated through several case studies. The results shows the improvements reached with the proposed strategies.

18. A NOVEL DISTANCE FOR CLUSTERING TO SUPPORT MIXED DATA ATTRIBUTES AND PROMOTE DATA RELIABILITY AND NETWORK LIFETIME IN LARGE SCALE WIRELESS SENSOR NETWORKS

Clustering based approaches in Wireless Sensor Networks helps in identifying the summarized data by exploiting the feature of data redundancy in sensor networks. Due to the inexpensive hardware used and unattended operation nature, nodes in the sensor networks are often prone to many failures malicious attacks and resource constraints and data collected in sensor networks are found to be unreliable. Moreover, the wide usages of sensor network in diverse application have put a constraint on sensor protocol to handle data of mixed types. To address the issues of energy

minimization and data reliability, we propose a distributed agglomerative cluster based anomaly detection algorithm termed DACAD to detect the faulty readings based on kNN approach. Additionally, to support applications with mixed data attributes, we design a heterogeneous distance function, HOEM to handle both continuous and nominal attributes. In this paper we have evaluated the performance of proposed algorithm in terms of false alarm rate, false positive rate and detection rate. Our results demonstrate that the proposed distance achieves a comparable detection rate with low false alarm rate with a significant reduction in computation and communication over head and operates with both continuous and nominal data.

19. CONCLUSION

An effort was taken to refer some of the recent leading journal articles related to Data Mining and its application areas. During this survey, it was found that Data Mining has almost played its role in most of the popular applications. It has helped to improve the performance of many organizations, do an effective survey on e-learning processes, improve the service quality of different organizations and Systems. This literature review would also help one to understand the different data mining techniques used in various fields.

REFERENCES

- [1] www.ijetse.com
- [2] www.ijcaonline.org
- [3] www.ejise.com
- [4] www.sciencedirect.com/science/article
- [5] www.researchconnector.net

AUTHORS' BIOGRAPHY



C.Grace Padma, Research Scholar in Computer Science & Applications, Manonmaniam Sundaranar University, Tirunelveli is now working as Associate Professor in MCA at RVS College of Arts & Science, Sulur, Coimbatore. Her interested areas of research include, Data Mining, Business Intelligence & Quality of service.



Dr.T.Arunkumar is a researcher, guide and an academician working on Mobile networks, spanning over MANETS, Wireless technology, involving QoS, location management, Handoff Issues and Security. He has published many articles in International & IEEE Journals. He has completed many Projects & Consultancy works.